数理解析研究所講究録
1043巻 1998年 207-214

n人囚人のジレンマにおける先見的安定性

東北大学経済学部 鈴木 明宏 (Akihiro Suzuki)
東京都立大学経済学部 武藤 滋夫 (Shigeo Muto)

1 はじめに

鈴木，武藤(1996)では囚人のジレンマに対して間接支配の下での安定集合を適用して分析を行った。本論文の目的はn人囚人のジレンマの1つである「湖の汚染」にChwe(1994)で提案された間接支配の下での安定集合を適用し、その結果について考察することである。構成は以下の通りである。第2節では湖の汚染のストーリーを述べそれを戦略形ゲームにモデル化する。第3節では間接支配と間接支配の安定集合を定義する。第4節，第5節は湖の汚染における安定集合がどうなるかを考察する。第4節では各プレイヤーが単独で動く場合を、また第5節では提携による行動を許す場合を扱う。第6節は本論文のまとめと今後の研究の方向性について述べる。

2 モデル

本論文で分析する「湖の汚染」は以下のようなストーリーである。湖の周りで操業しているn個の工場を考える。工場は生産のため湖から取水し、使用後汚染された水を排出する。各工場は汚水を浄化して排出するかどうかを決定する。浄化してから排出するにはbのコストがかかる。汚水のまま排出する場合にはコストはかからない。また、生産にはきれいな水を用いねばならず、汚染している工場数がkの場合取水時に浄化するためにkcのコストがかかる。これを戦略形ゲームで記述すると以下のようになる。

プレイヤーの集合は$N=\{1,2,\ldots,n\}$、プレイヤーiの戦略集合は$S_i=\{C,D\}$。ここでCとDはそれぞれ上のストーリーでの「浄化してから排出」と「浄化せず排出」に対応する。すると戦略の組は$x=(x_1,x_2,\ldots,x_n)\in\prod_{i\in N}S_i$で表される。ここで、全ての$i\in N$について$x_i=C$または$D$である。次に、$x$においてDをとるプレイヤー数を$ND(x)$、Cをとるプレイヤー数を$NC(x)\equiv n-ND(x)$で表すとプレイヤー$i$の利得は以下のようになる。

$$u_i(x)=\begin{cases}-b-cND(x) & \text{if } x_i=C\\ -cND(x) & \text{if } x_i=D\end{cases}$$

ここで$c<b<nc$とする。以下のことが成立することに注意する。

注意：① $x_i=D, y_i=C$で他のプレイヤーの戦略が等しいとき、$u_i(x)>u_i(y)$。
② $x_i=y_i$で$NC(x)>NC(y)$ならば$u_i(x)>u_i(y)$。

3 間接支配と安定集合

n人戦略形ゲーム$(N=\{1,2,\cdots,n\},(S_i),(u_i))$を考える。戦略の組$y$から提携$T\subset N$のメンバ

ーだけが動くことでxが達成されるとき$y \underset{T}{\to} x$で表す。この後の節では提携のサイズを1に制限することがあるので、$|T| \geq 2$のときには特に「共同して動く」ということにする。ここで$|T|$は集合Tの元の個数である。以下で定義されるのが Chwe(1994)において提案された間接支配である。

定義：以下の条件が成り立つときxはyを間接支配するといいx inddom yで表す。

戦略の組の列　$(y=)x^0, x^1, \ldots, x^m (=x)$　と提携の列　$T(1), T(2), \ldots, T(m)$　で

$x^{j-1} \to_{T(j)} x^j$, $u_i(x) > u_i(x^{j-1})$　$\forall i \in T(j)$　$(j=1,2,\ldots,m)$となるものが存在。

ここで$m=1$のとき特に、xはyを（直接）支配するということがありx dom yで表す。また、以下のことが成り立つ。

注意：x indom yのときx indom x^j（$j=1,\cdots,m-1$）も成立する

これを用いて安定集合が以下のように定義される。

定義：戦略の組の集合$V \subset \Pi S_i$が安定集合であるとは以下の2条件が成り立つことである。

（1）内部安定性：x indom x'となるような$x, x' \in V$は存在しない。

（2）外部安定性：任意の$x'' \notin V$に対してx indom x''となるような$x \in V$が存在する。

4　「湖の汚染」における安定集合（1）——各プレイヤーが単独でのみ動く場合——

はじめに各プレイヤーは単独でしか動くことができない、すなわち支配の定義において提携のサイズが1に制限される場合について考察する。ここでk_1をb/cを超える最小の自然数とする。すなわち、$(k_1-1)c < b < k_1 c$または$b=(k_1-1)c$となるような自然数k_1を考える。k_1は以下のような意味を持つ。ある戦略の組とそこで C をとるプレイヤーiを考える。すると、iを含んで(k_1-1)人までが C から D へ変更するとiの利得は利得が上昇し、k_1人以上になると利得が減少する。すると、xとyで D から C へ(k_1-1)人までが変化しているような2点x, yについてはx indom yが成り立つ。また、以下の補題も成立する。

補題1：$NC(x) \geq NC(y)$かつ　x indom yとなるような$x, y \in S$は存在しない。

証明：x indom yとすると、共同して動くことがないときの間接支配の定義より$x^{j-1} \to_{\{i(j)\}} x^j$, $u_{i(j)}(x) > u_{i(j)}(x^{j-1})$（$j=1,\cdots,m$）を満たすような$(y=)x^0,\cdots,x^m(=x)$が存在する。すると、間接支配の定義のところで述べた注意によりx dom x^{m-1}であるから、最後は誰かが C→D となって終わることになる。つまり、$NC(x^{m-1}) > NC(x)$。よって、$NC(x)=NC(y)$のとき補題が成り立つことを示せば十分である（図1参照）。そこで、$\tilde{x}$を$x^0,\cdots,x^{m-1}$の中で最後に$NC(x^k)=NC(x)$となる点とする。プレイヤーの集合を次の3つに分割する。

$$\text{C or D のままのプレイヤーの集合 } T_1 \equiv \{i \in N | \tilde{x}_i = x_i\}$$
$$\text{C→D と変わるプレイヤーの集合 } T_2 \equiv \{i \in N | \tilde{x}_i = C, x_i = D\}$$
$$\text{D→C と変わるプレイヤーの集合 } T_3 \equiv \{i \in N | \tilde{x}_i = D, x_i = C\}$$

$i \in T_1 \cup T_3$ は $u_i(\tilde{x}) \geq u_i(x)$ であるから動かない。よって $i \in T_2$ が動くしかないが、これは $\tilde{x}$ の定義に矛盾する。従って、x indom y となるような $x, y \in S$ は存在しない。　　　　　　（証明終）

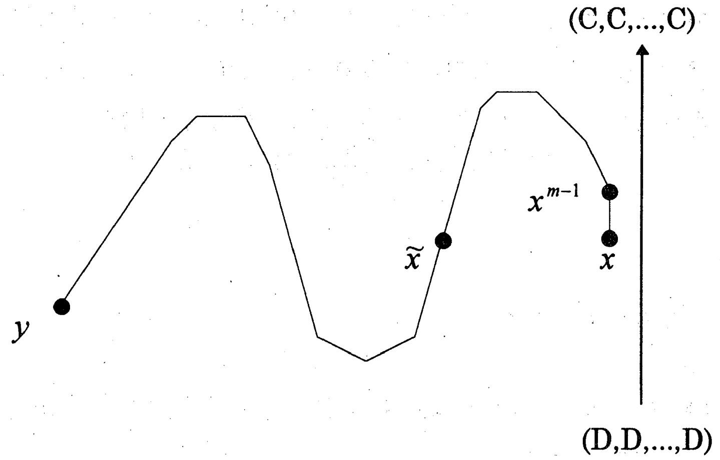

図１：x indom y とならないことの証明

以上の事実と補題１により、以下の定理が成り立つ。

定理１：各プレイヤーが単独でのみ動く場合の安定集合は

$$\{x \in \Pi S_i | NC(x) = k_1 a \ , a = 0,1,2,\ldots\}$$

が唯一存在する（図２参照）。

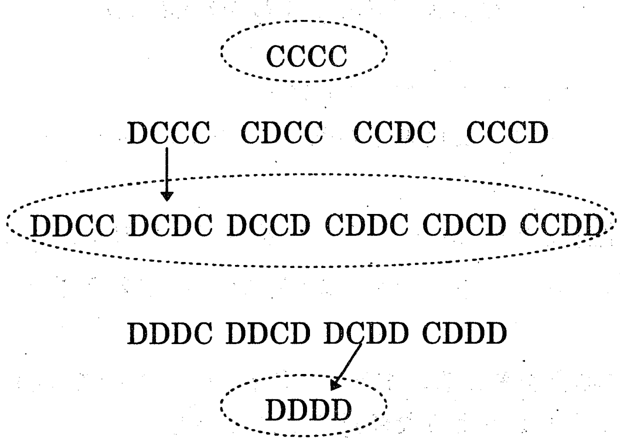

図２：安定集合の例（$N = 4, k_1 = 2$）

5　「湖の汚染」における安定集合（２）――共同した動きのある場合――

ここでは複数のプレイヤーが提携を形成し行動することが可能な場合を考える。ただし、一般

には外生的要因により全員提携が不可能な場合もありうる。その場合も分析できるように、提携可能なプレイヤー数がk_2に制限されるとする。もちろん、$k_2 = n$とすれば制限がない場合になる。ここでは$(k_1-1)c < b < k_1 c$となるか$b = (k_1-1)c$となるかで結果が少し変化するので場合分けを行うが証明は同様である。さらに形成可能な提携の規模に関して２つの場合に分かれる。

（１）$(k_1-1)c < b < k_1 c$

(I) $k_1 > k_2$の場合

「共同した動き」が導入されることで新たに生じる変化は、k_1人以上のプレイヤーが一緒に「D」から「C」に変更することで彼ら自身の利得が上昇するというものである。しかし、ここではその規模の提携は形成されないので、この場合には第４節と同様に

$$\left\{x \in \Pi S_i \,|\, NC(x) = k_1 a\,,\, a = 0,1,2,\ldots\right\}$$

が安定集合となる。

(II) $k_1 \le k_2$の場合

この場合、まず次の補題が成り立つ。

補題２：$NC(x) \ge \max\left\{k_1, n-k_1+1\right\}$となるような$x$１点のみからなる集合が安定集合で、それ以外には存在しない（図３参照）。

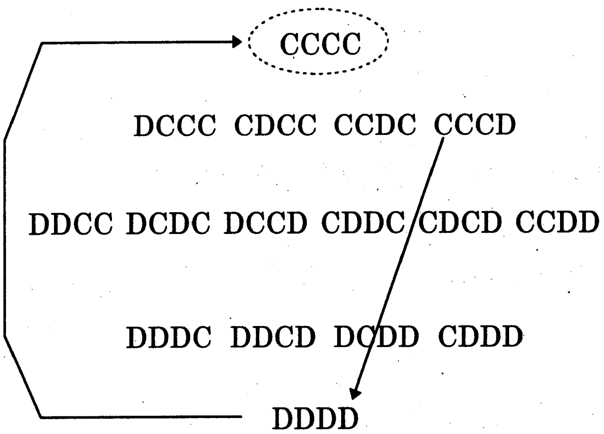

図３：安定集合の例（$n = 4, k_2 = 4$）

証明：まず、$\{x\}$が安定集合であることを示す。内部安定性が成り立つことは明らかであるから外部安定性のみ示せばよい。そこで以下の３つの場合について考える。

① $ND(x) < ND(y)$

まず、$y = (D,\ldots,D)$の場合を考える。すると、任意のiのついて$u_i(y) = -nc$。$NC(x) \ge k_1$，$b < k_1 c$であるから、$x_j = C$であるような任意のjについて

$$u_j(x) = -b - cND(x) \ge -b-(n-k_1)c > -nc = u_j(y)$$

であることに注意する。よって、$NC(x) \le k_2$なら$T(1) = \left\{i \in N | x_i = C\right\}$，$x^1 = x$とすればよい。$NC(x) > k_2$なら以下のように提携と戦略の組の列を構成する。まず、$l = NC(x) - k_2$，$N' = \left\{i \in N | x_i = C\right\}$とする。次に、$l$人の異なるプレイヤー$i(1), i(2), \cdots, i(l) \in N'$をとり、

$k=1,2,\cdots,l$ について $T(k)=\{i(k)\}$， $x^k_{i(k)}=C$ とする。さらに、

$$T(l+1)=N'-\big(T(1)\cup T(2)\cup\cdots\cup T(l)\big)$$

とする。すると明らかに、

$$(y=)x_0 \underset{T(1)}{\rightarrow} x_1 \underset{T(2)}{\rightarrow}\cdots \underset{T(l+1)}{\rightarrow} x\text{。}$$

$k=1,2,\cdots,l$ について、$x^k_i=D$ ならば $u_i(x^k)=-(n-k)c$。よって、$k=1,2,\cdots,l$ について、

$$\begin{aligned} u_{i(k)}(x)-u_{i(k)}(x^{k-1}) &= -b-(n-NC(x))c+(n-i)c \\ &\geq -b-(n-NC(x))c+(n+k_2-NC(x))c \\ &= -b+k_2c \geq -b+k_1c>0\text{。} \end{aligned}$$

また、全ての $i\in T(l+1)$ に対して $u_i(x)>u_i(x^l)$。従って、x は y を間接支配する。

次に、$y\neq(D,D,\cdots,D)$ の場合を考える。この場合、y から1人ずつ C から D に変えていき $(D,\ldots,D)$ に到達したら前の場合と同様にして構成する。これが間接支配の定義中の不等式条件を満たすことは、利得の定義のところで述べた注意②によりわかる。よってこの場合も x が y を間接支配する。

② $ND(x)=ND(y)$

$N'=\{i\in N|x_i=D,y_i=C\}$ とする。明らかに $i\in N'$ について $u_i(x)>u_i(y)$。それで、y から N' に属するプレイヤーが全員、一人ずつ C から D に変えていく。後は①の場合と同様である。。

③ $ND(x)>ND(y)$

もし $N'\equiv\{i\in N|x_i=D,y_i=C\}=\{i\in N|x_i\neq y_i\}$ ならば、そのようなプレイヤーが一人ずつ C から D に変えていけばよい。$N'\neq\{i\in N|x_i\neq y_i\}$ ならば、任意の $i\in N'$ に関して、

$$\begin{aligned} u_i(x)-u_i(y) &= -cND(x)+b+cND(y) \\ &> b-cND(x) \\ &\geq b-(k_1-1)c>0 \end{aligned}$$

である（$y\neq(C,C,\cdots,C)$ であるから式の2行目は厳密な不等号になる）。から x の方が利得が高くなる。そこで、D をとる人数が x と等しくなるまでそのようなプレイヤーが C から D へ変更する。その後は②と同様に考えればよい。

これまでの議論により $NC(x)\geq\max\{k_1,n-k_1+1\}$ であるような $\{x\}$ は安定集合であることがわかった。次に、このようなもの以外に安定集合が存在しないことを示す。もし存在するなら、それは $NC(x)<\max\{k_1,n-k_1+1\}$ となるような x からなるある集合 V である必要がある（今度は1点集合でなくともよい）。このような V は安定集合とはなり得ないことが以下のように示される。もし $k_1\geq n-k_1+1$ ならば、$NC(x)<k_1$。ところが、このような点 x は（$(D,D,\cdots,D)$ を除いて）第4節の結果より $(D,D,\cdots,D)$ に支配され、しかも $(D,D,\cdots,D)$ を支配できない。実際 $NC(x)<k_1$ ならば、$x_i=C$ となるような i は $y_i=D$ となるようなどんな y に対しても

$$u_i(x)<-b-(n-k_1)c<-nc\leq u_i(y)$$

であるので、i が D から C へ戦略を変更することはないからである。すると、$(D,D,\cdots,D)\in V$ でなければならないが、$(D,D,\cdots,D)$ は $(C,\ldots,C)$ を支配できないため、V は安定集合とはなり得ない。他方、もし $k_1<n-k_1+1$ ならば $NC(x)<n-k_1+1$、すなわち $ND(x)\geq k_1$。すると、

そのようなxについて

$$u_i(x) \leq -k_1 c < -b$$

が任意のiに対して成り立つからxは$(C,\ldots,C)$を支配できない。従って、いずれの場合にもVは安定集合とはなり得ない。 (証明終)

次に、$\max\{k_1, n-k_1+1\}$を解釈するが、実は$NC(x) \geq \max\{k_1, n-k_1+1\}$であることは$x$が厳密に個人合理的かつパレート効率的であることと等価であることが以下の補題で示される。ここでxが厳密に個人合理的であるとは$u_i(x) > \min_{x_{-i} \in \Pi_{j \neq i} S_j} \max_{x_i \in S_i} u_i(x_i, x_{-i})$が全ての$i$について成り立つことである。また、$x$がパレート効率的であるとは以下が成り立つような$y$が存在しないことである。全ての$i$について$u_i(y) \geq u_i(x)$、かつある$i$について$u_i(y) > u_i(x)$が成立する。

補題３：$(k_1-1)c < b < k_1 c$となるような自然数k_1が存在するとする。このとき以下のことが成立する。

(i) xが厳密に個人合理的な点（ここでは全ての$i \in N$に対して$u_i(x) > -nc$）であるための必要十分条件は$NC(x) \geq k_1$。

(ii) xがパレート効率的な点であるための必要十分条件は$NC(x) \geq n-k_1+1$。

証明：(i) $x_i = C$であるようなiについて考えれば十分。$u_i(x) = -b - c \cdot (n - NC(x))$であるから

$$NC(x) \geq k_1 \text{ならば} u_i(x) + nc > (NC(x) - k_1)c \geq 0、$$

$$NC(x) \leq k_1 - 1 \text{ならば} u_i(x) + nc < (NC(x) - (k_1-1))c \leq 0。$$

(ii) まず、$NC(x) < n - k_1 + 1$とすると、$ND(x) > k_1 - 1$。$y \equiv (C, C, \cdots, C)$とすると全ての$i \in N$について$u_i(y) = -b$。すると、

$$x_i = C \text{ならば} u_i(x) = -b - c \cdot ND(x) < -b、$$

$$x_i = D \text{ならば} u_i(x) = -c \cdot ND(x) < -(k_1-1)c < -b。$$

よって、xはパレート効率的ではない。

次に、xは$NC(x) \geq n - k_1 + 1$かつパレート効率的でないとする。すると、パレート効率的でないことより全ての$i \in N$に対して$u_i(y) > u_i(x)$となるような点$y \in \Pi S_i$が存在する。そこで次の３つの集合を定義する。

$$T_C \equiv \{i \in N | x_i = D, y_i = C\}$$

$$T_D \equiv \{i \in N | x_i = C, y_i = D\}$$

$$T \equiv T_C \cup T_D = \{i \in N | x_i \neq y_i\}$$

$NC(x) \geq n - k_1 + 1$すなわち$ND(x) \leq k_1 - 1$であるから$|T_C| \leq k_1 - 1$であることに注意する。$T_D = \varnothing$とする。すると$(k_1-1)c < b$であるから$|T_C| \geq k_1$となり矛盾。よって$T_D \neq \varnothing$。次に、$T_C = \varnothing$とすると$T = T_D$。$T = N$とすると$x = (C, C, \cdots, C)$，$y = (D, D, \cdots, D)$という可能性しかなく、全てのiについて$u_i(y) < u_i(x)$となり矛盾。よって$T \neq N$。すると、$i \in N - T$について$x_i = y_i$かつ$T = T_D$であるから、$u_i(y) < u_i(x)$となり矛盾。ゆえに$T_C \neq \varnothing$。すると、$i \in T_C$に対して

$$u_i(y) = u_i(x) - b + (|T_C| - |T_D|)c$$

$$\leq u_i(x) - b + (k_1 - 1)c$$
$$< u_i(x)$$

となりやはり矛盾。従って、x はパレート効率的である。　　　　　　　　　　　　（証明終）

（2） $b = (k_1 - 1)c$

この場合にも（1）とほぼ同様の結果が成り立つ。

(I) $k_1 > k_2$ ならば $\{x \in \Pi S_i \mid NC(x) = (k_1 - 1)a\ ,\ a = 0,1,2,\ldots\}$ が安定集合となる。

(II) $k_1 \leq k_2$ ならば $NC(x) \geq \max\{k_1, n - k_1 + 2\}$ となるような x 1点のみからなる集合が安定集合となる。ただし、この場合にはそれ以外にも安定集合が存在する可能性がある。例として図4をあげておく。

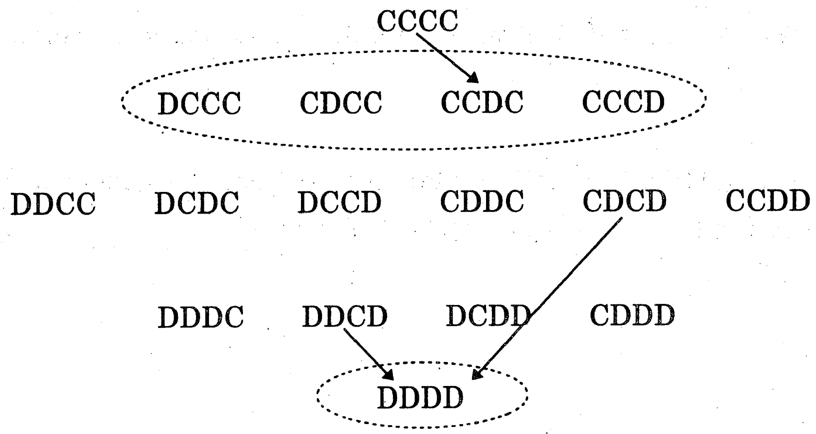

図4：安定集合の例（$n = 4, k_1 = k_2 = 4$）

これらの結果を要約すると以下の定理になる。

定理2：共同した動きがあるとする。もし $k_1 > k_2$ ならば、$(k_1 - 1)c < b < k_1 c$ である場合には

$$\{x \in \Pi S_i \mid NC(x) = k_1 a\ ,\ a = 0,1,2,\ldots\}$$

が、また $b = (k_1 - 1)c$ の場合には

$$\{x \in \Pi S_i \mid NC(x) = (k_1 - 1)a\ ,\ a = 0,1,2,\ldots\}$$

が唯一の間接支配の安定集合である。もし $k_1 \leq k_2$ ならば、厳密に個人合理的かつパレート効率的な1点よりなる集合は間接支配の安定集合である。さらに、$(k_1 - 1)c < b < k_1 c$ ならばそれ以外の安定集合は存在しない。

6　終わりに

この分析の結果、以下のことが導かれる。各プレイヤーがそれぞれ独立に行動する、といった状況では全ての工場が湖を汚染するというパレート非効率な状態が起こる可能性を避けられない。また、これは各プレイヤーが先を見通すことなく行動する、つまり直接支配に関する安定集合を考えたときにも安定集合は

$$\{x \in \Pi S_i \mid NC(x) = 2a\ ,\ a = 0,1,2,\ldots\}$$

となることが示され、やはり全員が湖を汚染する可能性が存在する。一方、ある程度先を見て行動する可能性（間接支配）と提携を組んで行動する可能性を導入（共同した動きの導入）することで湖の汚染がかなり解消されるという結果が得られた。

なお、ここで分析したモデルではプレイヤーに関して直面する状況が対称的である。またこの利得関数では、他のプレイヤーの状態に関わらず、戦略の変更が利得に与える影響の変化は常に一定である。そこで今後の方向性としては、より一般の利得関数や非対称生を導入した場合についての研究が考えられる。

参考文献

Chwe,M.S.-K. (1994),"Farsighted Coalitional Stability," Journal of Economic Theory 63,299-325.

鈴木 明宏, 武藤 滋夫 (1996), "Farsighted Stability in Prisoner's Dilemma," 京大数理解析研究所講究録 978 不確実性を含むシステムにおける最適化手法。